

Strategic Directions for the Program for Cooperative Cataloging

Regina Romano Reynolds
Library of Congress

The PCC and Continuing Resources

Continuing Resources in a Changing Landscape

John Comaromi's 1983 paper, "Cataloging Theory and Serials," begins:

"What is man? A question that man should not ask, I suppose. What are serials? Another question that man has asked—with equal success in receiving a durable answer."¹

The difficulty of defining serials in 1983 has only increased in the ensuing years of rapidly evolving online resources and library attempts to control these resources. One such attempt, revising AACR2 to accommodate seriality, gave us the new term "continuing resources," an overarching category that encompasses both serial and integrating modes of issuance. Continuing e-resources resources are constantly in flux, not only in content but in format, as publishers experiment with the best ways to provide their content, and as new technology becomes available. What are continuing resources becoming? How should we define and handle them? Perhaps Comaromi was right in suggesting we should not even ask these questions.

Regardless of what we call them, PCC should be mindful of the changing nature of continuing e-resources but also of the decline in print continuing resources. Stories continue to appear in news sources about long-standing print newspapers ceasing or becoming online only and print magazines such the 131-year-old *Ladies Home Journal* ceasing as a regular monthly. The recent *New York Times* "Innovation Report" recommends, in effect, the revolutionary overthrow of the primacy of the print *Times* in favor of its digital content being regarded as primary. Are we ready for these changes?

There is a great likelihood that e-resources will evolve much faster than cataloging rules and practices. One example of a relatively new resource that is coming into prominence for research is blogs. Scholarly blogs can contain the most current information or most recent discourse on a scientific or current events topic. Another type of resource of increased importance is the data set, with some journals requiring publication of an article's data set or its deposit into a repository such as DRYAD, a repository for evolutionary biology data. We also see that resources available on the web are being adapted for use on mobile devices, and now some resources are issued only for viewing on mobile devices. Institutional repositories are becoming valuable sources of open access articles, potentially replacing an institution's journal. The multitude of what we now regard as different formats or genres of online resources might be coalescing into less differentiated data streams to be consumed on various devices. How should these resources be handled in the MARC environment and in a future BIBFRAME/linked data environment?

We also need to consider the consequences of dividing resources into categories like “serial” and “integrating resource” using definitions that might be completely irrelevant to the researcher. Will categories like “serial” and “integrating resource” soon become obsolete? What resources will be missed in a library catalog or discovery system search when a user does not understand our terms or does not realize that an online directory will not appear in a search limited to books because it has been treated as a serial, or that a journal which has morphed into a web page might not appear in search results limited to “periodicals.” A recent ISSN request received in the U.S. ISSN Center involved a web site that later gave birth to a journal that “lived” on that web site for a time and then the journal was absorbed back into the web site. The content was later also made available as an ongoing database by a vendor. The same title has continued to be used for all of this content. How many resources are there? How many records are needed? And what are the consequences of the various ways we might treat situations like this? This example might seem extreme but in the volatile world of online resources, these kinds of changes happen all the time. How is the library to know about all of these ongoing changes even if we wanted to try to keep our cataloging current? How do we define “current?”

Why Catalog Continuing Resources?

Peter Burnhill, director of The Keepers Registry, a UK register of organizations that are archiving e-journals, calls the journal article, “the object of desire.” In other words, access to the journal is usually only needed in order to get access to the article. How important is it to have extensive descriptions of widely available commercial journals? It is already assumed, even by some librarians, that vendors like ProQuest are creating the catalog records they provide along with access to content, rather than what we know to be vendors’ heavy reliance on LC/CONSER records. If there were fewer LC/CONSER records, would vendors fill the gap adequately or would we be able to adjust our perceived needs and our dependent processes and make do?

In the years ahead, it will become increasingly important to be able to provide hard evidence to justify the need for the various aspects of the cataloging process, weighing the relative value of time spent on description, time spent on authority records, time spent on subject analysis, etc. Perhaps evidence will show that catalog records for many continuing resources are largely for internal identification, management, and control, thus allowing us to take a different approach to the creation of at least some records.

The Promise of Linked Data

PCC is well positioned to help assess the potential of linked data to make library metadata and thus discovery of library resources more pervasive on the web, and simultaneously more efficient and effective to provide. PCC is well positioned to help shape libraries’ linked data future. Linked data holds special promise for serials that should not be overlooked. One of the long-standing challenges of serials is how to handle their propensity to change titles. The debate over keeping all information on one bib record, despite changes of title (latest entry) vs. creating a new record for every major title change (successive entry) has raged for over 50 years. A more recent but equally contentious debate over whether to create single or multiple records for resources in multiple physical forms erupted when online resources first came to

the fore in the late 1980s and has not completely abated since. In the current record- and format-based systems, these challenges have seemed insurmountable. For serials, freeing the descriptive and access elements from the confines of the serial record has the potential to solve these problems, allowing appropriate elements to be combined or separated on the fly to provide the user with literally a picture of a serial's title and format history or to present exactly the title and format desired.

BIBFRAME, RDA/FRBR, Interoperability

As has become very clear, especially since the first BIBFRAME report, library metadata is shackled by MARC, confined in the "deep web," inaccessible to search engines and to other data that might link to library data, enriching and enlarging its effectiveness. BIBFRAME proposes to remedy this problem by using RDF triples, statements that operate in the Semantic Web. Given the challenges of modeling serials, it remains to be seen how continuing resources, especially serials with complex title histories and format versions, can fit into the BIBFRAME model.

BIBFRAME has simplified the FRBR's WEMI hierarchy to "work" and "instance," and can accommodate RDA's use of WEMI by using a profile to map WEMI elements into those belonging to BIBFRAME work and BIBFRAME instance. However, even this adaptation is problematic for serials. Echoing Comaromi's question, "what is a serial?" a new question, "What is a serial work?" has arisen and has yet to be fully answered for either FRBR/RDA or BIBFRAME. RDA simply accepted AACR2's treatment of major title changes as requiring a new description. While that acceptance happily kept RDA serial records aligned with AACR2 serial records, ISSN, and the entire serial control apparatus, the lack of agreement on what constitutes a serial work may impede mapping serials effectively in BIBFRAME. BIBFRAME's ability to accommodate serials is a critical aspect of BIBFRAME that has yet to be resolved but must be undertaken.

One problem of accommodating serials in a hierarchical model such as FRBR or even BIBFRAME is that serials frequently do not behave in a hierarchical or linear fashion, but in more varied branching patterns, including offshoots, splits and mergers. "Descendants" in a serials "family tree" do not always inherit properties from their "ancestors," frustrating WEMI relationships which presume this inheritance and thus put most relationships at the work or expression level rather than at the manifestation level where many serial relationships take place.

The ISSN International Centre, in collaboration with the Bibliothèque nationale de France, has developed PRESSoo, a project which has modeled serials in manner more consistent with their non-hierarchical nature. PRESSoo takes into account change over time, something also lacking in FRBR. Initial exchanges between BIBFRAME and PRESSoo have taken place and should continue. PRESSoo is likely to be adopted by the ISSN International Centre and perhaps other ISSN centers. There is also potential for ISSN centers besides the U.S. to use BIBFRAME (e.g., Germany, Scandinavian countries). BIBFRAME might benefit from adopting or adapting PRESSoo's approach to serials but at a minimum, BIBFRAME and PRESSoo must interoperate.

Interoperability with a broad range of other web-based systems and models will be critical to realizing BIBFRAME's stated goals. MARC's monopoly as a bibliographic data model will likely never be seen again. Various options to format bibliographic data are being explored in the U.S. and abroad in parallel with BIBFRAME. PCC should monitor BIBFRAME developments as well as monitor the developments in the broader community so as to help ensure the interoperability of BIBFRAME with other potential bibliographic data formats its members might use.

A final thought about BIBFRAME is that an editor or input application should be developed to handle batch processing as well as individual record production.

Batch Processing

There was a time when some speculated that cataloging workloads might stabilize despite increases in online continuing resources because equal numbers of print resources would cease. Although some print resources have ceased, print publishing does not seem to have diminished enough to compensate for the proliferation of online continuing resources, particularly online open access journals—even if some are of dubious quality. Identifying, describing, and controlling increasing numbers of resources with decreasing staff in an evolving digital environment calls for re-assessing current work processes.

Despite decades of automation, catalog records are generally hand-crafted one at a time, even when a group of titles from the same publisher and with much of the same metadata is in hand. Preliminary experimentation in the U.S. ISSN Center has demonstrated that reasonable efficiencies can be realized via a batch process that uses a spreadsheet mapped to MARC fields. As long as there is a need for separate descriptions for various formats of a resource, a similar batch record creation process might be used to create the separate records from a single input spreadsheet. Given the potential of the relatively primitive experimental process still in its early stages, one can only envision even greater efficiencies from a more sophisticated application.

Conclusion

The PCC has been a powerful and effective force in cataloging cooperation as well as standards development and application. Although it may be difficult for some to accept, PCC should communicate the hard reality that cataloging as we know it is at a crossroads: it needs to be re-assessed in light of the current library and web environment, human resource challenges, and the future linked data or other possible environments. We in cataloging need to make some evidence-based decisions about what to hold on to, emphasize, grow or develop—for example, our authoritative names, subject and work identification and the ability for our data to interoperate on the web, and what to let go-- perhaps detailed descriptions, especially for widely available commercial resources. We cannot be certain about the exact nature of our future technological or social environment but as computer scientist Alan Kay has said, "the best way to predict the future is to invent it."

¹ *Cataloging and Classification Quarterly*, issue 2-3 (1983), "AACR2 and Serials, the American View," p. 9.